An Extension of the Kalman Filter for a Class of Measurement Models Inspired by Wide-Baseline Stereo[1]

Talha Manzoor[2] and Abubakr Muhammad[3]

Abstract— **We consider the application of the Kalman Filter to systems with a certain type of measurement model. In addition to the state at the current time step, these measurements also depend on the state from one time step earlier. Although such measurement models are not encountered very often, they do appear in some practical control applications in robotic vision. In this paper, we derive a generalized Extended Kalman Filter from a modified form of the basic Bayes filter. The dependence of the measurement model on the previous state is made explicit through the law of total probability, which we include as an additional step in the standard prediction-correction cycle of the Bayes filter. We find that this dependence results in an increase in the measurement noise covariance. We present the mathematical formulation along with the proof of the filter equations. In the end, we demonstrate the filter through a simulation of Monocular SLAM with wide-baseline stereo measurements.**

I. Introduction

Filtering and estimation are pervasive tools in control applications requiring accurate state estimates in the presence of noisy sensors. The Kalman Filter [1] provides the Minimum Mean Squared Error (MMSE) estimate for linear systems, and is one of the most widely used techniques in this category. The Bayes filter algorithm (of which the Kalman filter is a special case) consists of a two step prediction-correction cycle. In this cycle, the error in the state propagated through the process model is corrected through measurements described by the measurement model. This measurement model describes the dependence of the measurement on the *present* state and possibly the control. We are concerned here with measurements which include an additional dependence on the state from one time step earlier.

The idea of measurements depending on an earlier state is somewhat counter-intuitive thus we spend some time here to describe our motivation. Our work is inspired by a particular type of measurement encountered in robotic vision, commonly known as motion-stereo or wide-baseline stereo [2]. Stereo vision is frequently employed in vision applications as a cheap and accurate sensor, consisting of two cameras simultaneously measuring the image projections of a single object in order to fully reconstruct its location in the environment. However, the length of the base-line separating the two cameras is a serious bottleneck in the ability to reconstruct points situated at large depths from the stereo pair [3]. Applications requiring depth measurements of distant scenes need an extremely large baseline which is not practically feasible especially if the cameras are to be mounted on a mobile robot [4]. Vision experts have tackled this problem through wide-baseline stereo where a single camera observes a static scene at two distinct points in time from viewpoints whose relative transformation is known [5]. The scene is then reconstructed by the conventional stereo vision algorithm where the stereo pair now consists of the same camera but at different time instants. Thus the disparity measurement now depends on the camera pose from one time step earlier and so, it is not covered by the measurement model assumed by the standard Kalman Filter. As an example application we consider motion-stereo measurements in Simultaneous Localization And Mapping [6] (SLAM), a problem studied extensively by the robotics community, which we introduce in detail in Section V.

In this paper we determine the effect of the inclusion of such measurement models in the Extended Kalman Filter framework. We accomplish this by deriving the EKF equations from the Bayes filter algorithm. The conventional Bayes filter is modified to include the dependence of the measurement model on the previous state through the law of total probability. The generalized EKF is then determined by following the outline of the standard Kalman filter derivation [7]. It is shown that the result is an increase in measurement noise covariance. This increase is different than what is obtained by augmenting the state at consecutive time steps into a single state vector. Thus this procedure is not equivalent to state augmentation. As one might expect, the increase is dependent on statistics of the process noise, marginal effect of the process noise on the state dynamics and marginal effect of the previous state on the current measurement.

The structure of the paper is as follows. Section II defines the problem mathematically. Section III presents the equations of the generalized EKF and Section IV proves their correctness through a mathematical derivation. Section V employs Monocular SLAM as an example application of the modified filter. Section VI describes the simulation environment and the results. We conclude in Section VII.

II. Problem Statement

We wish to apply the Kalman Filter to the following nonlinear discrete time system

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathrm{f}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k), \\ \mathbf{y}_{k+1} &= \mathrm{h}(\mathbf{x}_{k+1}, \mathbf{x}_k) + \mathbf{v}_{k+1}, \end{aligned} \tag{1}$$

[1] This work has been funded by a LUMS Faculty Initiative Fund (FIF) and the German Academic Exchange Service (DAAD) for the project RoPWat (Robotic Profiling for Clearing Watercourses).

[2]Talha Manzoor is a PhD student in Electrical Engineering at Lahore University of Management Sciences, Lahore, Pakistan 13060023 at lums.edu.pk

[3]Abubakr Muhammad is with the faculty of Electrical Engineering, Lahore University of Management Sciences, Lahore, Pakistan abubakr at lums.edu.pk

where $\mathbf{x}_k$ represents the n dimensional state vector at time step k, $\mathbf{u}_k$ is the control input, $\mathbf{w}_k$ is the process noise, $\mathbf{y}_k$ is the measurement vector, $\mathbf{v}_k$ is the measurement noise, $\mathrm{f}(\cdot,\cdot,\cdot)$ represents the process model and $\mathrm{h}(\cdot,\cdot)$ represents the measurement model. Furthermore, it is assumed that

$$\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k); \quad \mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k),$$

where $\mathbf{Q}_k$ and $\mathbf{R}_k$ are the covariance matrices of the process and measurement noise respectively. $\mathbf{w}_k$ and $\mathbf{v}_k$ are also assumed to be white and uncorrelated to each other such that

$$\mathrm{E}[\mathbf{w}_i, \mathbf{w}_j^T] = \mathrm{E}[\mathbf{v}_i, \mathbf{v}_j^T] = 0 \ \forall i \neq j; \quad \mathrm{E}[\mathbf{w}_i, \mathbf{v}_j^T] = 0 \ \forall i, j.$$

We further assume (for reasons to become apparent shortly), that the process model $\mathrm{f}(\cdot,\cdot,\cdot)$ is invertible i.e. given the state at time $k+1$, along with the control input and noise vectors at time k, it is possible to calculate the state at time k. We call this relation the *inverse process* and it is given as

$$\mathbf{x}_k = \mathrm{g}(\mathbf{x}_{k+1}, \mathbf{u}_k, \mathbf{w}_k).$$

Note that the system represented in (1) is different in the sense that the measurement model $\mathrm{h}(\cdot,\cdot)$ includes $\mathbf{x}_k$ the state vector at the previous time step, as an argument. We now proceed to generalizing the Extended Kalman Filter in order to include this dependence.

III. The Generalized EKF Equations

The Generalized EKF equations for the system in (1) are given as follows

Time Update Equations

$$\hat{\mathbf{x}}_{k+1|k} = \mathrm{f}\left(\hat{\mathbf{x}}_{k|k}, \mathbf{u}_k, 0\right), \tag{2}$$
$$\mathbf{P}_{k+1|k} = \mathbf{F}_{\mathbf{x}_k} \mathbf{P}_{k|k} \mathbf{F}_{\mathbf{x}_k}^T + \mathbf{F}_{\mathbf{w}_k} \mathbf{Q}_k \mathbf{F}_{\mathbf{w}_k}^T, \tag{3}$$

Measurement Update Equations

$$\mathbf{W}_{k+1} = \mathbf{R}_{k+1} + \mathbf{H}_{\mathbf{x}_k} \mathbf{G}_{\mathbf{w}_k} \mathbf{Q}_k \mathbf{G}_{\mathbf{w}_k}^T \mathbf{H}_{\mathbf{x}_k}^T, \tag{4}$$
$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k} \mathbf{H}_{\mathbf{x}_{k+1}}^T \Big(\mathbf{H}_{\mathbf{x}_{k+1}} \mathbf{P}_{k+1|k} \mathbf{H}_{\mathbf{x}_{k+1}}^T + \mathbf{W}_{k+1}\Big)^{-1}, \tag{5}$$
$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}\big(\mathbf{y}_{k+1} - \mathrm{h}\left(\hat{\mathbf{x}}_{k+1|k}, \hat{\mathbf{x}}_{k|k}\right)\big), \tag{6}$$
$$\mathbf{P}_{k+1|k+1} = \left(\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{\mathbf{x}_{k+1}}\right)\mathbf{P}_{k+1|k}, \tag{7}$$

where $\hat{\mathbf{x}}_{k+1|k}$ and $\mathbf{P}_{k+1|k}$ are the *a priori* state estimate and covariance, and $\hat{\mathbf{x}}_{k+1|k+1}$ and $\mathbf{P}_{k+1|k+1}$ are the *a posteriori* state estimate and covariance respectively. $\mathbf{F}_{\mathbf{x}_k}$, $\mathbf{F}_{\mathbf{w}_k}$, $\mathbf{G}_{\mathbf{w}_k}$, $\mathbf{H}_{\mathbf{x}_{k+1}}$ and $\mathbf{H}_{\mathbf{x}_k}$ are the Jacobian matrices of the process, inverse process and the measurement models w.r.t. the subscripted vectors respectively. $\mathbf{K}_k$ represents the Kalman gain at time k. We call $\mathbf{W}_k$ the *measurement covariance matrix* as it captures the uncertainty in measurement, both due to the measurement noise and the dependence of measurement on the previous state. It can be seen from (4) that $\mathbf{W}_k \succ 0, \forall k$ and thus it is a valid covariance matrix. In fact as we will see in Section IV it is the covariance of the distribution associated with the expected measurement.

Note here that the difference between the standard EKF and the EKF given by (2) – (7) lies in the calculation in (4). If there is no dependence of the measurement model on the previous state then $\mathbf{W}_k = \mathbf{R}_k$ and the filter reduces to the standard EKF. However if this dependence does exist, then it results in an increase in the *effective* measurement noise covariance. As this increase is due to the uncertainty associated with the state at the previous time step, one might expect this increase to be dependent on the process noise, the marginal contribution of the process noise to the inverse process, and the marginal contribution of the previous state to the current measurement. This is exactly what (4) represents. We now present the mathematical derivation of this filter.

IV. Mathematical Derivation

Here we present the proof of our Generalized EKF of Section III. Our proof is inspired by that of the standard EKF given in [7]. Our starting point is the basic Bayes filter algorithm which consists of the following two steps

1) Compute the a priori state distribution through the law of total probability using the process model and a posteriori distribution of the previous time step.
2) Compute the a posteriori distribution through the Bayes rule using the measurement model and the a priori distribution computed in step 1.

We extend this filter by explicitly representing the dependence of the measurement on the previous state through a second application of the law of total probability. This modified form of the Bayes filter is given as follows

The Modified Bayes Filter

Step 1: state prediction

$$\overline{\mathrm{bel}}(\mathbf{x}_{k+1}) = \mathrm{p}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k}, \mathbf{u}_k) = \int \mathrm{p}(\mathbf{x}_{k+1}|\mathbf{x}_k, \mathbf{y}_{0:k}, \mathbf{u}_k)\mathrm{p}(\mathbf{x}_k|\mathbf{y}_{0:k}, \mathbf{u}_k)\mathrm{d}\mathbf{x}_k,$$

Step 2: measurement prediction

$$\mathrm{p}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1}, \mathbf{y}_{0:k}, \mathbf{u}_k) = \int \mathrm{p}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1}, \mathbf{y}_{0:k}, \mathbf{u}_k, \mathbf{x}_k)\,\mathrm{p}(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{y}_{0:k}, \mathbf{u}_k)\mathrm{d}\mathbf{x}_k,$$

Step 3: state correction

$$\mathrm{bel}(\mathbf{x}_{k+1}) = \mathrm{p}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k+1}, \mathbf{u}_k) = \eta\,\mathrm{p}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1}, \mathbf{y}_{0:k}, \mathbf{u}_k)\mathrm{p}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k}, \mathbf{u}_k),$$

where the function $\overline{\mathrm{bel}}(\mathbf{x}_k)$ represents the prior *belief* of the state $\mathbf{x}_k$, $\mathrm{bel}(\mathbf{x}_k)$ represents the posterior belief of $\mathbf{x}_k$ and $\mathbf{y}_{0:k}$ represents all measurement vectors from time 0 to k.

A. Part 1: State Prediction

We begin with the first step of the modified Bayes filter. The terms appearing on the R.H.S of Step 1 are normally distributed with mean and covariance given as

$$\underbrace{\mathrm{p}(\mathbf{x}_{k+1}|\mathbf{x}_k,\mathbf{y}_{0:k},\mathbf{u}_k)}_{\sim\mathcal{N}\left(\mathbf{x}_{k+1};\,\mathrm{f}(\hat{\mathbf{x}}_k,\mathbf{u}_k,0)+\mathbf{F}_{\mathbf{x}_k}(\mathbf{x}_k-\hat{\mathbf{x}}_k);\,\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_kF_{\mathbf{w}_k}^T\right)},$$

and,

$$\underbrace{\mathrm{p}(\mathbf{x}_k|\mathbf{y}_{0:k},\mathbf{u}_k)}_{\sim\mathcal{N}(\mathbf{x}_k;\hat{\mathbf{x}}_k;\mathbf{P}_k)},$$

where $\mathcal{N}(x;\mu;\sigma^2)$ represents the Gaussian distribution for random variable x with mean μ and variance σ^2. Note that from here onwards we will take $\hat{\mathbf{x}}_k$ to mean $\hat{\mathbf{x}}_{k|k}$ (similarly for $\mathbf{P}_k$). We will now prove that $\overline{\mathrm{bel}}(\mathbf{x}_{k+1})$ is in fact a Gaussian random vector with mean and covariance given by (2) and (3) respectively. After substituting the PDF's of $\mathrm{p}(\mathbf{x}_{k+1}|\mathbf{x}_k,\mathbf{y}_{0:k},\mathbf{u}_k)$ and $\mathrm{p}(\mathbf{x}_k|\mathbf{y}_{0:k},\mathbf{u}_k)$ in Step 1 of the modified Bayes filter, $\overline{\mathrm{bel}}(\mathbf{x}_{k+1})$ can be written as,

$$\overline{\mathrm{bel}}(\mathbf{x}_{k+1})=\eta\int\exp\{-\mathbf{L}_{k+1}\}\mathrm{d}\mathbf{x}_k,$$

where

$$\begin{aligned}\mathbf{L}_{k+1}&=\frac{1}{2}(\mathbf{x}_k-\hat{\mathbf{x}}_k)^T{\mathbf{P}_k}^{-1}(\mathbf{x}_k-\hat{\mathbf{x}}_k)\\&+\frac{1}{2}\Big(\big(\mathbf{x}_{k+1}-\mathrm{f}(\hat{\mathbf{x}}_k,\mathbf{u}_k,0)-\mathbf{F}_{\mathbf{x}_k}(\mathbf{x}_k-\hat{\mathbf{x}}_k)\big)^T(\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T)^{-1}\\&\qquad\big(\mathbf{x}_{k+1}-\mathrm{f}(\hat{\mathbf{x}}_k,\mathbf{u}_k,0)-\mathbf{F}_{\mathbf{x}_k}(\mathbf{x}_k-\hat{\mathbf{x}}_k)\big)\Big).\end{aligned}$$

To solve the integral in closed form, we decompose $\mathbf{L}_{k+1}$ as $\mathbf{L}_{k+1}=\mathbf{L}(\mathbf{x}_{k,+1}\mathbf{x}_k)+\mathbf{L}(\mathbf{x}_{k+1})$. This decomposition enables us to move $\mathbf{L}(\mathbf{x}_{k+1})$ outside the integral, simplifying it to

$$\overline{\mathrm{bel}}(\mathbf{x}_{k+1})=\eta\exp\{-\mathbf{L}(\mathbf{x}_{k+1})\}\int\exp\{-\mathbf{L}(\mathbf{x}_{k+1},\mathbf{x}_k)\}\mathrm{d}\mathbf{x}_k.$$

Also, we choose $\mathbf{L}(\mathbf{x}_{k+1},\mathbf{x}_k)$ such that the value of the integral does not depend upon $\mathbf{x}_{k+1}$ and hence, can be submerged within the constant η. We now describe how such a function may be obtained.

Note that $\mathbf{L}_{k+1}$ is quadratic in $\mathbf{x}_k$. Thus it is possible to take $\mathbf{L}(\mathbf{x}_{k+1},\mathbf{x}_k)$ as a quadratic function in $\mathbf{x}_k$ as well. If this is so, then it may be interpreted as the function appearing in the exponential of a Gaussian PDF (which always integrates to a constant). The mean and variance will be given by the minimum and inverse of the curvature of $\mathbf{L}_{k+1}$ (w.r.t $\mathbf{x}_k$) as $\min\limits_{\mathbf{x}_k}\mathbf{L}(\mathbf{x}_{k+1},\mathbf{x}_k)$ and $\left\{\dfrac{\partial^2\mathbf{L}(\mathbf{x}_{k+1},\mathbf{x}_k)}{\partial\mathbf{x}_k^2}\right\}^{-1}$. And so, the resulting distribution will be defined entirely through $\mathbf{L}(\mathbf{x}_{k+1})$ as $\overline{\mathrm{bel}}(\mathbf{x}_{k+1})=\eta\exp\{-\mathbf{L}(\mathbf{x}_{k+1})\}$. Through this procedure we get the required function as

$$\begin{aligned}&\mathbf{L}(\mathbf{x}_{k+1},\mathbf{x}_k)=\\&\frac{1}{2}\Big[\mathbf{x}_k-\boldsymbol{\Psi}_k\mathbf{F}_{\mathbf{x}_k}^T\Big(\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T\Big)^{-1}\Big(\mathbf{x}_{k+1}-\mathrm{f}(\mathbf{x}_k,\mathbf{u}_k,0)\Big)\\&\qquad+\boldsymbol{\Psi}_k\Big(\mathbf{F}_{\mathbf{x}_k}^T(\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T)^{-1}\mathbf{F}_{\mathbf{x}_k}\hat{\mathbf{x}}_k-\mathbf{P}_k^{-1}\hat{\mathbf{x}}_k\Big)\Big]^T\\&\boldsymbol{\Psi}_k^{-1}\Big[\mathbf{x}_k-\boldsymbol{\Psi}_k\mathbf{F}_{\mathbf{x}_k}^T\Big(\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T\Big)^{-1}\Big(\mathbf{x}_{k+1}-\mathrm{f}(\mathbf{x}_k,\mathbf{u}_k,0)\Big)\\&\qquad+\boldsymbol{\Psi}_k\Big(\mathbf{F}_{\mathbf{x}_k}^T(\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T)^{-1}\mathbf{F}_{\mathbf{x}_k}\hat{\mathbf{x}}_k-\mathbf{P}_k^{-1}\hat{\mathbf{x}}_k\Big)\Big],\end{aligned}$$

where

$$\boldsymbol{\Psi}_k=\Big(\mathbf{F}_{\mathbf{x}_k}^T(\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T)^{-1}\mathbf{F}_{\mathbf{x}_k}+{\mathbf{P}_k}^{-1}\Big)^{-1}.$$

Finally, $\mathbf{L}(\mathbf{x}_{k+1})$ can be obtained by subtracting $\mathbf{L}(\mathbf{x}_{k+1},\mathbf{x}_k)$ from $\mathbf{L}_{k+1}$. The large size of the expression makes it infeasible to display here and we direct the interested reader to [8] for the full calculation. We find that upon computing this difference, all terms containing $\mathbf{x}_k$ vanish as required. Furthermore, $\mathbf{L}(\mathbf{x}_{k+1})$ turns out to be quadratic in $\mathbf{x}_{k+1}$ which means that it is a valid function for the exponential of a Gaussian PDF. The mean and covariance of the resulting PDF are found as

$$\hat{\mathbf{x}}_{k+1|k}=\min_{\mathbf{x}_{k+1}}\mathbf{L}(\mathbf{x}_{k+1});\quad\mathbf{P}_{k+1|k}=\left\{\frac{\partial^2\mathbf{L}(\mathbf{x}_{k+1})}{\partial\mathbf{x}_{k+1}^2}\right\}^{-1}.$$

Carrying out this calculation gives us the following result

$$\begin{aligned}\hat{\mathbf{x}}_{k+1|k}&=\mathrm{f}(\mathbf{x}_k,\mathbf{u}_k,0),\\\mathbf{P}_{k+1|k}&=\mathbf{F}_{\mathbf{x}_k}\mathbf{P}_k\mathbf{F}_{\mathbf{x}_k}^T+\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T.\end{aligned}$$

Thus $\overline{\mathrm{bel}}(\mathbf{x}_{k+1})$ is indeed Gaussian with mean $\hat{\mathbf{x}}_{k+1|k}$ and covariance $\mathbf{P}_{k+1|k}$ both of which are given above. This proves the correctness of (2) and (3).

B. Part 2: Measurement Prediction

The procedure we follow here is exactly the same as that followed in Section IV-A. We begin with Step 2 of the modified Bayes filter, the terms appearing on the R.H.S of which, are normally distributed with mean and covariance given as

$$\underbrace{\mathrm{p}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1},\mathbf{y}_{0:k},\mathbf{u}_k,\mathbf{x}_k)}_{\sim\mathcal{N}\left(\mathbf{y}_{k+1};\;\begin{matrix}\mathrm{h}(\hat{\mathbf{x}}_{k+1},\hat{\mathbf{x}}_k)+\mathbf{H}_{\mathbf{x}_{k+1}}(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1})\\+\mathbf{H}_{\mathbf{x}_k}(\mathbf{x}_k-\hat{\mathbf{x}}_k)\end{matrix}\;;\mathbf{R}_k\right)},$$

and,

$$\underbrace{\mathrm{p}(\mathbf{x}_k|\mathbf{x}_{k+1},\mathbf{y}_{0:k},\mathbf{u}_k)}_{\sim\mathcal{N}\left(\mathbf{x}_k;\;\begin{matrix}\mathrm{g}(\hat{\mathbf{x}}_{k+1},\mathbf{u}_k,0)+\\\mathbf{G}_{\mathbf{x}_{k+1}}(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1})\end{matrix}\;;\mathbf{G}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{G}_{\mathbf{w}_k}^T\right)}.$$

After making these substitutions the integral can be expressed as $\mathrm{p}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1},\mathbf{y}_{0:k},\mathbf{u}_k)=\int\exp\{-\mathbf{M}_{k+1}\}\mathrm{d}\mathbf{x}_k$, where

$$\begin{aligned}&\mathbf{M}_{k+1}=\\&\frac{1}{2}\Big(\mathbf{y}_{k+1}-\mathrm{h}(\hat{\mathbf{x}}_{k+1},\hat{\mathbf{x}}_k)-\mathbf{H}_{\mathbf{x}_{k+1}}(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1})-\mathbf{H}_{\mathbf{x}_k}(\mathbf{x}_k-\hat{\mathbf{x}}_k)\Big)^T\\&R_{k+1}^{-1}\Big(\mathbf{y}_{k+1}-\mathrm{h}(\hat{\mathbf{x}}_{k+1},\hat{\mathbf{x}}_k)-\mathbf{H}_{\mathbf{x}_{k+1}}(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1})-\mathbf{H}_{\mathbf{x}_k}(\mathbf{x}_k-\hat{\mathbf{x}}_k)\Big)\\&+\frac{1}{2}\Big(\mathbf{x}_k-\mathrm{g}(\hat{\mathbf{x}}_{k+1},\mathbf{u}_k,0)-\mathbf{G}_{\mathbf{x}_{k+1}}(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1})\Big)^T\\&\Big(\mathbf{G}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{G}_{\mathbf{w}_k}^T\Big)^{-1}\Big(\mathbf{x}_k-\mathrm{g}(\hat{\mathbf{x}}_{k+1},\mathbf{u}_k,0)-\mathbf{G}_{\mathbf{x}_{k+1}}(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1})\Big).\end{aligned}$$

Just as before, we decompose $\mathbf{M}_{k+1}$ as $\mathbf{M}_{k+1}=\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1},\mathbf{x}_k)+\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1})$. This decomposition allows the following simplification

$$\begin{aligned}&\mathrm{p}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1},\mathbf{y}_{0:k},\mathbf{u}_k)=\\&\quad\eta\exp\{-\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1})\}\int\exp\{-\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1},\mathbf{x}_k)\}\mathrm{d}\mathbf{x}_k.\end{aligned}$$

Again, we choose $\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1},\mathbf{x}_k)$ such that the integrand becomes a valid Gaussian PDF (upto a normalizing constant) and so, integrates to a fixed value regardless of $\hat{\mathbf{x}}_k$. We obtain this function by calculating the minimum and inverse of the curvature of $\mathbf{M}_{k+1}$ (w.r.t $\mathbf{x}_k$) and so, the resulting distribution will be defined entirely through $\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1})$ as $\mathrm{p}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1},\mathbf{y}_{0:k},\mathbf{u}_k)=\eta\exp\{-\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1})\}$. $\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1})$ is found through the following expression

$$\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1})=\mathbf{M}_{k+1}-\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1},\mathbf{x}_k).$$

Once we calculate this difference, we find that all terms including $\mathbf{x}_k$ vanish as required. The final mean and covariance of the resulting distribution are calculated through the minimum and curvature of $\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1})$ as

$$\hat{\mathbf{y}}_{k+1}=\min_{\mathbf{x}_{k+1}}\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1});\ \mathbf{W}_{k+1}=\left\{\frac{\partial^2\mathbf{M}(\mathbf{y}_{k+1},\mathbf{x}_{k+1})}{\partial\mathbf{x}_{k+1}^2}\right\}^{-1}.$$

This gives us

$$\hat{\mathbf{y}}_{k+1}=\mathrm{h}(\hat{x}_{k+1},\hat{\mathbf{x}}_k)+\mathbf{H}_{\mathbf{x}_{k+1}}(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1}),$$

and,

$$\mathbf{W}_{k+1}=\mathbf{R}_{k+1}+\mathbf{H}_{\mathbf{x}_k}\mathbf{G}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{G}_{\mathbf{w}_k}^T\mathbf{H}_{\mathbf{x}_k}^T.$$

This proves the correctness of (4).

C. Part3: Correction / Update

We begin with Step 3 of the modified Bayes filter. The terms on the RHS are normally distributed as

$$\underbrace{\mathrm{p}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1},\mathbf{y}_{0:k},\mathbf{u}_k)}_{\sim\mathcal{N}\left(\mathbf{y}_{k+1};\mathrm{h}(\hat{\mathbf{x}}_{k+1},\hat{\mathbf{x}}_k)+\mathbf{H}_{\mathbf{x}_{k+1}}(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1});\mathbf{W}_{k+1}\right)},$$

and,

$$\underbrace{\mathrm{p}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k},\mathbf{u}_k)}_{\sim\mathcal{N}\left(\mathbf{x}_{k+1};\mathrm{f}(\hat{\mathbf{x}}_k,\mathbf{u}_k,0);\mathbf{F}_{\mathbf{x}_k}\mathbf{P}_k\mathbf{F}_{\mathbf{x}_k}^T+\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T\right)}.$$

As easily observed, the product is given by the exponential $\mathrm{bel}(\mathbf{x}_{k+1})=\eta\exp\{-\mathbf{J}_{k+1}\}$, where

$$\begin{aligned}\mathbf{J}_{k+1}=\frac{1}{2}&\Big(\mathbf{y}_{k+1}-\mathrm{h}(\hat{\mathbf{x}}_{k+1},\hat{\mathbf{x}}_k)-\mathbf{H}_{\mathbf{x}_{k+1}}(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1})\Big)^T\\&\Big(\mathbf{R}_{k+1}+\mathbf{H}_{\mathbf{x}_k}(\mathbf{G}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{G}_{\mathbf{w}_k}^T)\mathbf{H}_{\mathbf{x}_k}^T\Big)^{-1}\\&\Big(\mathbf{y}_{k+1}-\mathrm{h}(\hat{\mathbf{x}}_{k+1},\hat{\mathbf{x}}_k)-\mathbf{H}_{\mathbf{x}_{k+1}}(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1})\Big)\\&+\frac{1}{2}\Big(\mathbf{x}_{k+1}-\mathrm{f}(\hat{\mathbf{x}}_k,\mathbf{u}_k,0)\Big)^T\\&\Big(\mathbf{F}_{\mathbf{x}_k}\mathbf{P}_k\mathbf{F}_{\mathbf{x}_k}^T+\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T\Big)^{-1}\Big(\mathbf{x}_{k+1}-\mathrm{f}(\hat{\mathbf{x}}_k,\mathbf{u}_k,0)\Big).\end{aligned}$$

This function is quadratic in $\mathbf{x}_{k+1}$ and hence it represents a Gaussian distribution. The mean and covariance of this distribution is found as usual through the first two derivatives with respect to $\mathbf{x}_{k+1}$ as

$$\hat{\mathbf{x}}_{k+1}=\min_{\mathbf{x}_{k+1}}\mathbf{J}_{k+1};\ \text{and}\ \mathbf{P}_{k+1}=\left\{\frac{\partial^2\mathbf{J}_{k+1}}{\partial\mathbf{x}_{k+1}^2}\right\}^{-1}.$$

The covariance turns out to be given by

$$\begin{aligned}\mathbf{P}_{k+1}^{-1}=&\Big(\mathbf{F}_{\mathbf{x}_k}\mathbf{P}_k\mathbf{F}_{\mathbf{x}_k}^T+\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T\Big)^{-1}+\\&\mathbf{H}_{\mathbf{x}_{k+1}}^T\Big(\mathbf{R}_{k+1}+\mathbf{H}_{\mathbf{x}_k}(\mathbf{G}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{G}_{\mathbf{w}_k}^T)\mathbf{H}_{\mathbf{x}_k}^T\Big)^{-1}\mathbf{H}_{\mathbf{x}_{k+1}}.\end{aligned}\tag{8}$$

Next, putting the first derivative equal to zero and solving for $\mathbf{x}_{k+1}$, we get the mean as

$$\begin{aligned}\hat{\mathbf{x}}_{k+1}=&\hat{\mathbf{x}}_{k+1|k}\\&+\mathbf{P}_{k+1}\mathbf{H}_{\mathbf{x}_{k+1}}^T\Big(\mathbf{R}_{k+1}+\mathbf{H}_{\mathbf{x}_k}(\mathbf{G}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{G}_{\mathbf{w}_k}^T)\mathbf{H}_{\mathbf{x}_k}^T\Big)^{-1}\\&\Big(\mathbf{y}_{k+1}-\mathrm{h}(\hat{\mathbf{x}}_{k+1},\hat{\mathbf{x}}_k)\Big).\end{aligned}$$

Defining the Kalman Gain as

$$\mathbf{K}_{k+1}=\mathbf{P}_{k+1}\mathbf{H}_{\mathbf{x}_{k+1}}^T\Big(\mathbf{R}_{k+1}+\mathbf{H}_{\mathbf{x}_k}(\mathbf{G}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{G}_{\mathbf{w}_k}^T)\mathbf{H}_{\mathbf{x}_k}^T\Big)^{-1},$$

we get

$$\hat{\mathbf{x}}_{k+1}=\hat{\mathbf{x}}_{k+1|k}+\mathbf{K}_{k+1}\big(\mathbf{y}_{k+1}-\mathrm{h}\left(\hat{\mathbf{x}}_{k+1|k},\hat{\mathbf{x}}_k\right)\big).$$

This proves the correctness of (6).

The Kalman Gain as given above, is a function of $\mathbf{P}_{k+1}$ which is at odds with the fact that we utilize $\mathbf{K}_{k+1}$ to calculate $\mathbf{P}_{k+1}$ in (7). To avoid this, we transform our current expression of $\mathbf{K}_{k+1}$, to include covariances other than $\mathbf{P}_{k+1}$. This transformation is done by multiplying the RHS by a cleverly chosen expression and its inverse as follows

$$\begin{aligned}\mathbf{K}_{k+1}=&\mathbf{P}_{k+1}\mathbf{H}_{\mathbf{x}_{k+1}}^T\Big(\mathbf{R}_{k+1}+\mathbf{H}_{\mathbf{x}_k}(\mathbf{G}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{G}_{\mathbf{w}_k}^T)\mathbf{H}_{\mathbf{x}_k}^T\Big)^{-1}\\&\Big(\mathbf{H}_{\mathbf{x}_{k+1}}(\mathbf{F}_{\mathbf{x}_k}\mathbf{P}_k\mathbf{F}_{\mathbf{x}_k}^T+\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T)\mathbf{H}_{\mathbf{x}_{k+1}}^T\\&+(\mathbf{R}_{k+1}+\mathbf{H}_{\mathbf{x}_k}(\mathbf{G}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{G}_{\mathbf{w}_k}^T)\mathbf{H}_{\mathbf{x}_k}^T)\Big)\\&\Big(\mathbf{H}_{\mathbf{x}_{k+1}}(\mathbf{F}_{\mathbf{x}_k}\mathbf{P}_k\mathbf{F}_{\mathbf{x}_k}^T+\mathbf{F}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{F}_{\mathbf{w}_k}^T)\mathbf{H}_{\mathbf{x}_{k+1}}^T\\&+(\mathbf{R}_{k+1}+\mathbf{H}_{\mathbf{x}_k}(\mathbf{G}_{\mathbf{w}_k}\mathbf{Q}_k\mathbf{G}_{\mathbf{w}_k}^T)\mathbf{H}_{\mathbf{x}_k}^T)\Big)^{-1}.\end{aligned}$$

Multiplying the terms and carrying out a simplification similar to that of [7], we get the following expression for the Kalman Gain

$$\mathbf{K}_{k+1}=\mathbf{P}_{k+1|k}\mathbf{H}_{\mathbf{x}_{k+1}}^T\left(\mathbf{H}_{\mathbf{x}_{k+1}}\mathbf{P}_{k+1|k}\mathbf{H}_{\mathbf{x}_{k+1}}^T+\mathbf{W}_{k+1}\right)^{-1}$$

This proves the correctness of (5).

The expression in (8) for $\mathbf{P}_{k+1}$ that we have calculated involves an inversion which is a heavy computation for high dimensional state spaces. To avoid this, we carry out a transformation to express $\mathbf{P}_{k+1}$ in a more convenient fashion. Recognizing certain terms in (8) and substituting for them, gives us

$$\mathbf{P}_{k+1}=\left(\mathbf{P}_{k+1|k}^{-1}+\mathbf{H}_{\mathbf{x}_{k+1}}^T\mathbf{W}_{k+1}^{-1}\mathbf{H}_{\mathbf{x}_{k+1}}\right)^{-1}.$$

Applying the *matrix inversion lemma* [9] to this expression, we get

$$\begin{aligned}\mathbf{P}_{k+1}=&\mathbf{P}_{k+1|k}\\&-\mathbf{P}_{k+1|k}\mathbf{H}_{\mathbf{x}_{k+1}}^T\Big(\mathbf{W}_{k+1}+\mathbf{H}_{\mathbf{x}_{k+1}}\mathbf{P}_{k+1|k}\mathbf{H}_{\mathbf{x}_{k+1}}^T\Big)^{-1}\mathbf{H}_{\mathbf{x}_{k+1}}\mathbf{P}_{k+1|k}.\end{aligned}$$

Further simplifying

$$\mathbf{P}_{k+1} = \mathbf{P}_{k+1|k} - \mathbf{K}_{k+1}\mathbf{H}_{\mathbf{x}_{k+1}}\mathbf{P}_{k+1|k},$$
$$= \left(\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{\mathbf{x}_{k+1}}\right)\mathbf{P}_{k+1|k},$$

which proves the correctness of (7).

V. Example Application: Wide-Baseline Stereo in Monocular SLAM

As an application we consider here the problem depicted in Figure 1: A camera navigating in a 2-D environment, taking measurements of point features. The task is to track the positions of the camera itself and the unknown point features in the environment. This problem is commonly known in the robotics community as Simultaneous Localization And Mapping (SLAM) and has been studied for a range of different sensors. We are interested here in the case of a single camera (or Monocular SLAM). It has recently been pointed out that it might be beneficial to incorporate the disparity measurement in addition to the conventional projective camera measurements [8]. The measurement equations of disparity however, include the state variables at the previous time step and so, our extension to the standard Kalman filter applies here. For the particular setting depicted in Figure 1, we take the state to be the following vector consisting of the camera pose and map variables

$$\mathbf{x}_k = \left[\begin{array}{cccccccc} {}^{c}x_k & {}^{c}x'_k & {}^{c}y_k & {}^{c}y'_k & {}^{c}\theta_k & {}^{c}\theta'_k & {}^{L2}x_k & {}^{L2}y_k \end{array}\right]^T,$$

where ${}^{c}x_k, {}^{c}y_k$ and ${}^{c}\theta_k$ are the location and orientation variables of the camera at time k, and ${}^{c}x'_k, {}^{c}y'_k$ and ${}^{c}\theta'_k$ are their respective velocities. ${}^{L2}x_k$ and ${}^{L2}y_k$ represent the location of landmark 2. Note that as depicted in Figure 1, we categorize landmarks into two types. The first type of landmarks are unknown landmarks. The location of such landmarks is to be estimated by the EKF and so, is included in the state vector. The second type of landmarks referred to as *beacons*, are landmarks whose location is precisely known beforehand and are placed in the environment for the sole purpose of maintaining system *observability* [10].

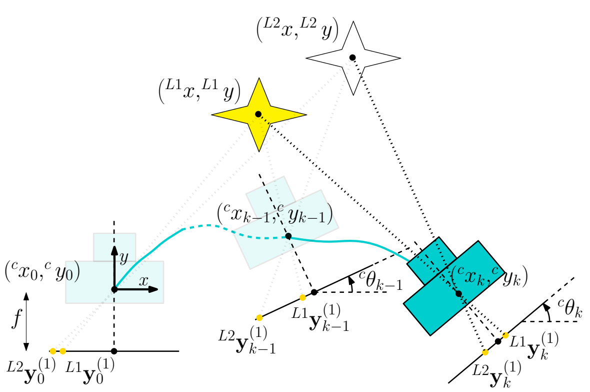

Fig. 1: Monocular state sequence in 2-D. The yellow star represents a beacon whereas the white star represents an unknown landmark included in the state vector.

For the state dynamics, we assume a constant velocity motion model [11] given as

$$\mathbf{x}_{k+1} = \begin{pmatrix} \bar{\mathbf{F}} & 0 & 0 & 0 \\ 0 & \bar{\mathbf{F}} & 0 & 0 \\ 0 & 0 & \bar{\mathbf{F}} & 0 \\ 0 & 0 & 0 & \mathbf{I} \end{pmatrix} \mathbf{x}_k + \mathbf{w}_k,$$

where $\bar{\mathbf{F}} = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix}$ with Δt being the time step. The process noise covariance is given by

$$\mathbf{Q}_k = \begin{pmatrix} q_v\bar{\mathbf{Q}} & 0 & 0 & 0 \\ 0 & q_v\bar{\mathbf{Q}} & 0 & 0 \\ 0 & 0 & q_\theta\bar{\mathbf{Q}} & 0 \\ 0 & 0 & 0 & q_l\mathbf{I} \end{pmatrix}$$

where $\bar{\mathbf{Q}} = \begin{pmatrix} \frac{1}{3}\Delta t^3 & \frac{1}{2}\Delta t^2 \\ \frac{1}{2}\Delta t^2 & \Delta t \end{pmatrix}$ and q_v, q_θ and q_l are the noise variances in the linear velocity, angular velocity and landmark position respectively.

The measurement vector is given by $\mathbf{y}_k = \left[\, {}^{L1}\mathbf{y}_k^T \; {}^{L2}\mathbf{y}_k^T \,\right]^T$ where ${}^{Li}\mathbf{y}_k$ is a 2×1 vector corresponding to landmark i. The first entry is the projective measurement and is given as

$${}^{Li}\mathbf{y}_k^{(1)} = f\frac{({}^{Li}x_k - \mathbf{x}_k^{(1)})\cos(\mathbf{x}_k^{(5)}) + ({}^{Li}y_k - \mathbf{x}_k^{(3)})\sin(\mathbf{x}_k^{(5)})}{-({}^{Li}x_k - \mathbf{x}_k^{(1)})\sin(\mathbf{x}_k^{(5)}) + ({}^{Li}y_k - \mathbf{x}_k^{(3)})\cos(\mathbf{x}_k^{(5)})},$$

where f is the focal length of the camera, $({}^{Li}x_k, {}^{Li}y_k) = ({}^{Li}x, {}^{Li}y)$ for $i=1$ and $({}^{Li}x_k, {}^{Li}y_k) = (\mathbf{x}_k^{(7)}, \mathbf{x}_k^{(8)})$ for $i=2$. The second entry is the disparity measurement and is given as

$${}^{Li}\mathbf{y}_k^{(2)} = \frac{f\left(\mathbf{x}_k^{(1)} - \mathbf{x}_{k-1}^{(1)}\right)^2 + f\left(\mathbf{x}_k^{(3)} - \mathbf{x}_{k-1}^{(3)}\right)^2}{\left(\mathbf{x}_k^{(1)} - \mathbf{x}_{k-1}^{(1)}\right)\left(\mathbf{x}_{k-1}^{(3)} - {}^{Li}y_k\right) - \left(\mathbf{x}_{k-1}^{(1)} - {}^{Li}x_k\right)\left(\mathbf{x}_k^{(3)} - \mathbf{x}_{k-1}^{(3)}\right)}.$$

It is this second component of the measurement vector due to which there is a dependence of the measurement model on the previous state $\mathbf{x}_{k-1}$. Further details of the measurement equations can be found in [8].

VI. Simulation

A. Setup

The simulation shown in this section has been implemented in MATLAB. In the scenario presented here, the robot follows a rhombus shaped trajectory with smoothed corners as shown in Figure 2. Velocity has been kept higher at the sides and lower at the corners. The focal length of the camera has been set at 20 pixel units. The process and measurement models are as given in Section V. The process model is a constant velocity one with process noise modeled as zero-mean Gaussian with variances of $10\,(\mathrm{m/s^2})^2$, $2\ \mathrm{deg}^2$ and $0.02\ \mathrm{m}^2$ for the linear velocities, angular velocities and landmark positions respectively. Measurement noise variables have also been drawn from zero-mean Gaussian distributions with variances of $0.074\ \mathrm{pixels}^2$ for projective and $74\ \mathrm{pixels}^2$ for disparity measurements. Length of one time step has been kept at 0.1 seconds after which the robot receives projective and disparity measurements of the landmarks and beacons. As seen in Figure 2 an incorrect initial estimate is provided with large uncertainty.

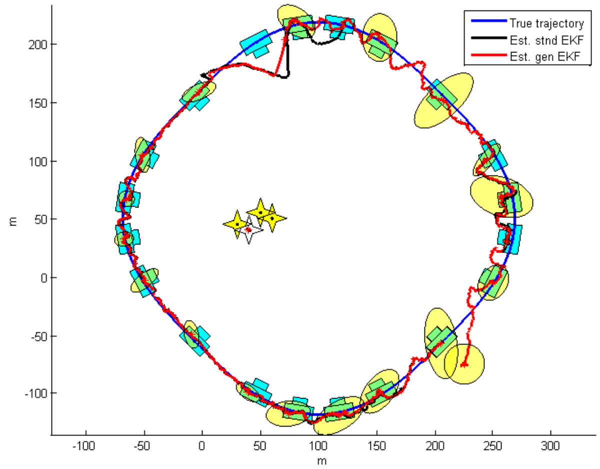

Fig. 2: A robot performing SLAM in 2-D. The cameras depict the true pose of the robot at selected time steps, equally distributed in time. The beacons are drawn in yellow whereas the estimated landmark is in white. Estimated trajectories for the standard and generalized EKF are shown. The ellipses represent the 95% confidence regions. An incorrect estimate with large uncertainty is provided.

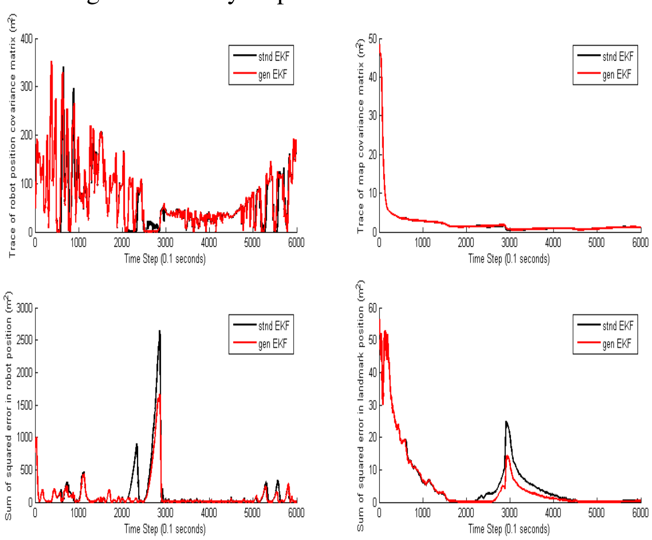

Fig. 3: Performance comparison between the conventional EKF and the generalized EKF presented in Section III. From top-left in counter clockwise order: 1) Trace of covariance matrix for robot location, 2) SSD for robot location, 3) SSD for landmark location, 4) Trace of covariance matrix for landmark location.

B. Results

The results of the simulation can be seen in Figures 2 & 3. Figure 2 shows tracking of the states in context of the environment. Due to incorrect initialization, there is a little inconsistency for some time after the filter is initialized. Overall, the filter shows consistent behavior. This can also be seen for the uncertainty ellipse of the estimated landmark. Figure 3 compares the performance between the conventional EKF and the modified EKF of Section III. The first two graphs show the trace of the position and map covariance matrices, whereas the last two show the sum of squared error (or Sum of Squared Difference (SSD)) between the true states and the estimated ones.

It is important to note here that the performance of the generalized EKF appears to be better simply because of the increased measurement noise covariance. For linear systems, the increase will be constant and thus this performance can be matched simply by increasing the measurement noise covariance in the standard EKF framework. The magnitude of this increase may be arrived at intelligently through the methodology proposed in this paper, or simply by tuning the noise parameters of the filter through empirical hit and trial.

VII. Conclusion

In this paper we have considered the application of the EKF to systems with measurement models dependent on states from one time step earlier. Through a modified form of the Bayes Filter, we have shown that this dependence amounts simply to an increase in the measurement noise covariance. For linear systems, this increase may be knowingly calculated offline and incorporated into the standard EKF, or tuned empirically. The derivation done here gives a theoretical justification for this increase, which may not be apparent simply by tuning the noise covariances. For non-linear systems however, this increase must be incorporated in the standard EKF equations at each time step, in the form we have derived here. Although in this paper the consideration of such measurement models has been inspired by a particular application in robotics, the work shown here may be applied to similar models encountered in other fields as well.

References

[1] G. Welch and G. Bishop, "An introduction to the kalman filter," 1995.

[2] J. Huber and V. Graefe, "Motion stereo for mobile robots," in *Industrial Electronics, 1993. Conference Proceedings, ISIE'93-Budapest., IEEE International Symposium on*. IEEE, 1993, pp. 263–270.

[3] R. Szeliski, *Computer vision: algorithms and applications*. Springer, 2010.

[4] C. F. Olson and H. Abi-Rached, "Wide-baseline stereo vision for terrain mapping," *Machine Vision and Applications*, vol. 21, no. 5, pp. 713–725, 2010.

[5] P. Pritchett and A. Zisserman, "Wide baseline stereo matching," in *Computer Vision, 1998. Sixth International Conference on*. IEEE, 1998, pp. 754–760.

[6] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part i," *Robotics & Automation Magazine, IEEE*, vol. 13, no. 2, pp. 99–110, 2006.

[7] S. Thrun, W. Burgard, D. Fox, *et al.*, *Probabilistic robotics*. MIT press Cambridge, MA, 2005, vol. 1.

[8] T. Manzoor, "Observability and disparity in monocular slam: A study," Master's thesis, Lahore University of Management Sciences, 2013.

[9] D. Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[10] F. A. Belo, P. Salaris, and A. Bicchi, "3 known landmarks are enough for solving planar bearing slam and fully reconstruct unknown inputs," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 2539–2545.

[11] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.